Cell, Vol. 80, 115–125, January 13, 1995, Copyright © 1995 by Cell Press

Hypermutation Generating the Sheep Immunoglobulin Repertoire Is an Antigen-Independent Process

Claude-Agnès Reynaud,* Corinne Garcia,* Wayne R. Hein,† and Jean-Claude Weill*
*Institut National de la Santé et de la Recherche Médicale
Unité 373
Institut Necker
Faculté de Médecine Necker–Enfants Malades
Université Paris 5
F-75730 Paris Cédex 15
France
†Basel Institute for Immunology
Grenzacherstrasse 487
CH-4002 Basel
Switzerland

Summary

Somatic hypermutation of light chain V genes during development of B cells in sheep ileal Peyer's patches was studied in three experimental conditions: in sterile fragments of the ileum surgically isolated from the gut during fetal life, in germ-free sheep, and in animals thymectomized during early fetal life. The somatic mutation pattern was found identical to control tissues in all three experiments. The same age-dependent amount of mutations, a higher than theoretical R/S ratio in complementarity-determining regions (CDRs), and a similar clustering of mutations in CDRs were observed. The mechanism, as estimated from the silent mutation pattern, appears to target mutations to CDRs; moreover, the major Vλ genes have a specific codon usage with a high purine content at the first two bases of the codons and a low content at the third position, which, together with a specific targeting of mutations to purines, favors replacement mutations in CDRs.

Introduction

The notion that B cells could undergo a stage of negative selection as they are produced in primary organs was proposed (Lederberg, 1959) and further demonstrated many years ago (Lawton and Cooper, 1974). It has also been known for a long time that immature B cells are prone to tolerogenic signals through their immunoglobulin receptor (Nossal and Pike, 1975; Metcalf and Klinman, 1976). More recently, it was clearly demonstrated in transgenic mouse models that high affinity recognition of a persistent antigen by newly formed B cells leads to their silencing in the bone marrow (Goodnow et al., 1988; Nemazee and Bürki, 1989). However, one cannot estimate at this moment what the contribution of negative selection is during normal B lymphopoiesis.

As B cells are generated in the bone marrow, they proceed from a pro-B to a pre-B and a B cellular phenotype (Rolink and Melchers, 1991), but the question of whether there are crucial developmental checkpoints involving interactions between surface immunoglobulin molecules and specific ligands at each one of these stages is still open at the moment. Moreover, based on the skewed expression of some V gene families during the transition between bone marrow and spleen, it was suggested that newly generated B cells could also be positively selected through their immunoglobulin receptors by internal or external antigens while they migrate to the periphery and that these interactions could shut off the recombinase machinery and direct B cell clones to the short- or long-lived peripheral pool (Bos and Meeuwsen, 1989; Gu et al., 1991; Viale et al., 1992; Ma et al., 1992; Huetz et al., 1993; Deenen and Kroese, 1993). However, this remains speculative, and despite the resemblance between the differentiation pathway of pre-T and pre-B cells, a clear experimental model supporting the concept of positive selection has not yet been produced for B cell development.

The generation of the sheep B cell compartment relies chiefly on production of B cells in ileal Peyer's patches (IPPs) during the second half of gestation and up to 6 months after birth (Reynolds and Morris, 1983a). We have more recently shown that sheep, like chickens, generate their preimmune repertoire in IPPs by a postrearrangement process acting on a fixed population of surface immunoglobulin M (IgM) B cell precursors (Reynaud et al., 1991). In sheep, like in chickens, these B cell precursors colonize gut-associated lymphoid tissue (GALT) follicles during fetal life and, thereafter, proliferate extensively within these follicles. However, contrary to chicken, the diversification process taking place in sheep B cells is not gene conversion, which has also been shown to be the major generator of B cell diversity in rabbit (Becker and Knight, 1990), but rather is ongoing hypermutation.

Our first analysis on normal IPP lymphoid tissue at different stages of development indicated that the hypermutation process starts before birth in the sterile environment of the fetus and then continues for several months after birth. Moreover, the accumulation of replacement mutations in complementarity-determining regions (CDRs), together with the high rate of cell death in situ (95%; Reynolds, 1986), seemed to imply that a selection was acting on newly formed B cells before their emigration to the periphery (Reynaud et al., 1991).

At birth, gut tissues are directly in contact with thousands of food and bacterial antigens. Moreover, specific epithelial cells are involved in capturing these antigens and putting them in contact with the neighboring lymphoid tissues (Landsverk, 1987). We therefore thought that sheep IPPs could represent an ideal model to test whether foreign antigens can positively select newly formed B cells through their immunoglobulin receptor and thus shape the pattern of a primary antibody repertoire.

Paradoxically, we show here that the accumulation of mutations over time on rearranged Vλ genes is independent of the presence of external antigens. Moreover, the major type of mutations observed and the specific nucleotide composition of the most utilized Vλ genes show that

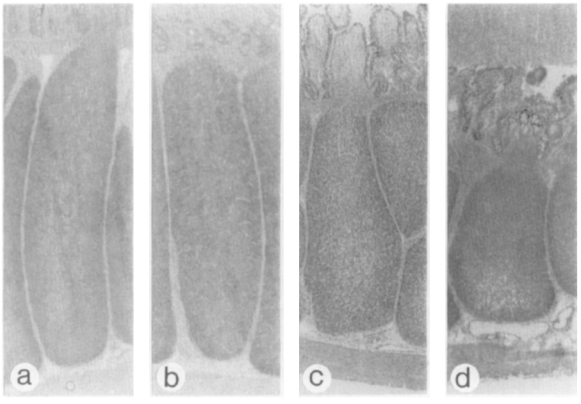

Figure 1. Histological Appearance of IPP Follicles

Representative frozen sections of Peyer's patches are shown to illustrate the relative sizes of individual follicles observed in different experimental conditions. (a) IPP follicle from a normal 41-day-old lamb. (b) IPP follicle from a 33-day-old thymectomized lamb. (c) Follicle from the ileum of a 37-day-old germ-free lamb. (d) Follicle from a sterile, isolated segment of ileum in a 41-day-old lamb (Nocht's azure eosin, x50).

the hypermutation process per se can target replacement mutations in CDRs.

Results

The rearrangement pattern of the λ light chain locus is restricted in sheep IPP B cells. After EcoRI digestion, two major rearranged fragments, which we named "type I" (around 20 kb) and "type II" (around 12 kb), can be observed (Reynaud et al., 1991). When these rearranged sequences are analyzed, they can be attributed unambiguously to specific germline genes. Therefore, we can follow the diversification of one or the other of these two rearrangements in various experimental conditions without resorting to any experimental cellular or immunological selection procedure.

Hypermutation Starts before Birth

We have shown previously that after polymerase chain reaction (PCR) amplification of rearranged fetal Vλ sequences there was an average of 0.9 mutations per sequence at approximately 1 week before birth (day 139), taking into account the error rate of the Taq polymerase (Reynaud et al., 1991). To obtain a more direct estimate of the extent of mutation occurring before birth without the incidence of polymerase errors introduced experimentally, we isolated seven genomic clones containing rearranged type I Vλ sequences at the same stage (144 days in utero). The sequences obtained confirmed that fetal rearranged sequences are mutated: nine mutations (plus one P sequence; Lafaille et al., 1989) in seven VJ sequences, with five of them in the V part and four in the J segment (data not shown).

The Hypermutation Process Remains Antigen Independent after Birth

We have set up two experimental conditions in which IPP follicles remain in a sterile environment during development.

Ileal Loops

A segment of IPP was isolated from gut-associated antigens by transsection of the terminal ileum at 117 days (experiments 1 and 3) and 118 days (experiment 2) of fetal life. The ileum was reanastomosed, and the isolated segment (ileal loop) was maintained in situ after ligation at both ends, with its nervous T lymphatic and vascular connections. The lambs were killed, and the isolated segments were sampled at 143 days of fetal life (experiment 1, "fetal loop"), 41 days (experiment 2, "6-week loop"), and 60 days (experiment 3, "8-week loop") after birth. In the 6-week and 8-week experiments, a segment of normal IPP adjacent to the operated fragment was taken and used as control tissue for molecular studies.

Phenotypic Analysis

When assessed at the histological level before birth, follicles in isolated ileal loops showed no differences in size as compared with normal ileum, and their cellular composition appeared similar. In contrast, at 6 and 8 weeks after birth, the follicles in the loop segment were less than half their normal size and appeared to be less densely populated with B cells (Figure 1). However, fluorescence-activated cell sorting (FACS) analysis showed that the cellular composition of the two types of IPPs from postnatal animals was virtually indistinguishable: both contained 1% (or less) T cells and 95%–98% of cells expressed B lineage–specific markers. The single consistent phenotypic difference detected by FACS analysis was a slightly higher level of expression of the CD45RA isotype of the leucocyte common antigen on lymphocytes from normal IPP (97%) as compared with isolated loop IPP (91%). In both control and loop IPP, immunohistology revealed the presence of low numbers of all T cell subsets ($CD4^+$, $CD8^+$, $\gamma\delta$ TCR^+) in the small interfollicular areas, consistent with the pattern expected in normal IPP tissue (Hein et al., 1989). The expression of B cell markers within follicles, the relative number and distribution of tingible body macrophages, and the pattern of dendritic cell staining appeared similar between normal and loop IPP.

Microbiology of the Loop

In one experiment, the lumen of the ileal loop was examined for its bacteriological content. It was found to be sterile: no bacteria or mycoplasmas were detected after culture under aerobic, anaerobic, and microaerophilic conditions. Normal ileum in the same animal contained about 20×10^6 bacteria per gram of lumen content. They were mainly Escherichia coli, Streptoccocus spp, and Clostridium perfringens, which are normal bacterial inhabitants of lamb guts.

Molecular Studies

In each experiment, IPP DNA was fractionated by gel electrophoresis after EcoRI digestion. The size region of rearrangement fragment I was eluted, the VJ sequences were cloned after PCR amplification, and rearranged sequences were determined. The sequences obtained were compared with the most homologous genomic sequence, Vλ 5.1, which differs at four positions from the type I rearranged sequences (Reynaud et al., 1991). There were

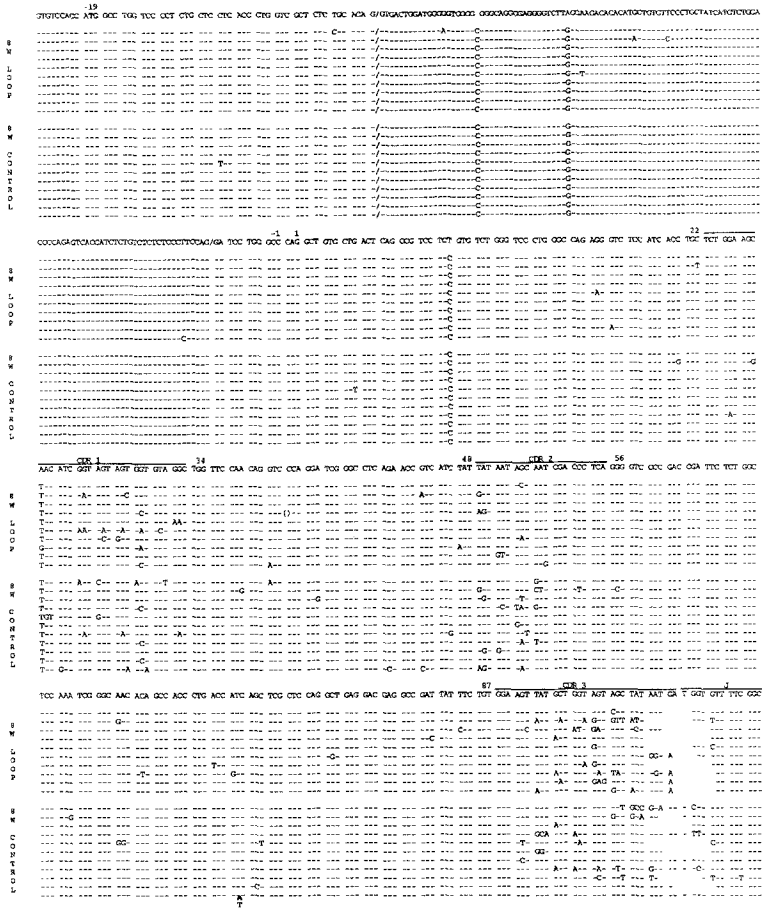

Figure 2. Somatic Mutations in Rearranged Vλ Genes of B Cells from Sterile Isolated Ileal Fragments

Rearrangement type I was isolated after gel fractionation of EcoRI-digested DNA from isolated ileal fragments at 8 weeks of age (8 W LOOP) and from normal adjacent tissue (8 W CONTROL) and was cloned after PCR amplification. Ten loop and eleven control sequences are compared with the 5.1 gene.

no quantitative or qualitative differences in the sequences obtained from the control tissue when compared with sequences from the loop fragments (Figure 2 and Table 1). They have a similar number of mutations per sequence (0.8 for the fetal loop, 4.3 and 7.8 for the 6-week and 8-week loops, respectively). They have a targeting of mutations in CDRs and a higher than theoretical R/S ratio for the CDRs. For the framework regions (FRs), the R/S ratio of the 8-week loop is higher than the theoretical value, but several replacement mutations are conservative changes (e.g., four Val↔Ile exchanges), and therefore, this value is not indicative of any specific effect provided by the loop configuration.

Germ-Free Animals

To exclude the possibility that some gut-associated bacterial antigens could be transported to the loop IPP through the circulation, we studied IPPs from gnotobiotic sheep. It was difficult to maintain lambs uncontaminated after 4–5 weeks. One animal that remained uncontaminated (7L, 37 days) and one that was killed 2 days after contamination (10L, 35 days) were studied.

Phenotypic Analysis

In lambs 7L and 10L, IPP follicles were approximately two-thirds the size of those observed in normal control animals, although their degree of cellularity seemed similar (Figure 1). There was a clear diminution of $CD4^+$ and $CD8^+$ T cells in the interfollicular areas and intestinal mucosa of these animals, but no obvious difference between the expression of B cell, dentritic cell, and tingible body macrophage markers as compared with normal animals. The jejunal Peyer's patches of these animals were also smaller in size and contained relatively fewer T cells.

Molecular Studies

As in the previous experiment, IPP DNA from lamb 7L was fractionated by gel electrophoresis, but the size region of rearrangement fragment II, which corresponds to the genomic 16.1 sequence, was eluted. The rearranged sequences analyzed were distributed in two allelic forms of this gene (Figure 3 and Table 1). On average, there were 5.2 mutations per sequence again with a strong clustering in CDRs and more replacement substitutions than the theoretical value, whereas there was a selection against replacement mutations in FRs. The same study on lamb 10L, which was killed at day 35, gave similar results (nine sequences corresponding to another closely related gene, 16/5, were analyzed; Table 1).

The Hypermutation Process Is Not Altered in Thymectomized Sheep

One lamb was thymectomized at 70 days of gestation and was killed at 33 days after birth. Such a thymectomy induces a severe T cell depletion during several weeks after birth (Hein et al., 1990), and we wanted to find out whether

Table 1. Somatic Mutations in Sheep Vσ Sequences in Loop, thymectomy, and Germ-Free Experiments

	Number of Sequences	5'nc	L	Intron	CDR Total	R	S	R/S	FR Total	R	S	R/S	J	Total Mutation	Total Corrected[e]	Mutation per Sequence	Clustering Ratio CDR/FR[f]
Fetal loop (5.1)	11	—	1	—	6	5	1	**5.0**	4	3	1	**3**	3	14	8.5	**0.8**	**3.8**
6-week loop (5.1)	10[a]	—	—	5	39	38	1	**38.0**	2	1	1	**1**	1	48	43	**4.3**	**49.9**
6-week control (5.1)	12	—	—	—	49	45	4	**11.3**	7	4	3	**1.3**	3	59	53	**4.4**	**17.9**
8-week loop (5.1)	10	—	1	5	61	56	5	**11.2**	14[b]	11	2	**5.5**	2	83	78	**7.8**	**11.2**
8-week control (5.1)	11	—	1	—	65	58	7	**8.3**	15[b]	8	6	**1.3**	7	88	82.5	**7.5**	**11.1**
5-week thymetcomy (5.1)	8	—	—	2	26	24	2	**12.0**	8	5	3	**1.7**	1	37	33	**4.1**	**8.3**
5-week thymectomy (16.1)	6	—	—	2	22	22	—	**>22[c]**	4	3	1	**3**	—	28	25	**4.2**	**13.1**
5-week control (5.1)	11[a]	—	—	2	45	45	—	**>45[c]**	10	5	5	**1.0**	2	59	53.5	**4.9**	**11.5**
5-week control (16.1)	7	—	—	0	27	23	4	**5.8**	6	3	3	**1.0**	3	35	31.5	**4.7**	**10.7**
5-week germ-free 7L (16.1)	15[a]	1	1	2	64	55	9	**6.1**	17	12	5	**2.4**	1	86	78.5	**5.2**	**9.0**
5-week germ-free 10L (16/5[d])	9	—	—	6[b]	28	25	3	**8.3**	8	5	3	**1.7**	1	43	38.5	**4.3**	**8.3**
Total	110	1	4	24	432	396	36	**11.0**	95[b]	60	33	**1.8**	24	580			**11.4**
Total 5.1					291	271	20	**13.6**									
Total 16.1					113	100	13	**7.7**									

[a] Includes one out-of-frame sequence.
[b] Includes deletions or insertions.
[c] >X stands for R/S for which R = X and S = 0.
[d] 16/5 is a new VL gene, with CDR1 of the 16.1 gene and CDR2 and CDR3 of 5.1.
[e] The error rate of the Taq polymerase (1.1×10^{-3}/bp in our experimental conditions; Reynaud et al., 1991) was deduced.
[f] Expressed as the ratio of mutation frequency per nucleotide of CDR and FR (length ratio of FR/CDR = 2.56 for the 5.1 gene, 2.38 for the 16.1 gene, and 2.5 on the average for the sequences tabulated above.

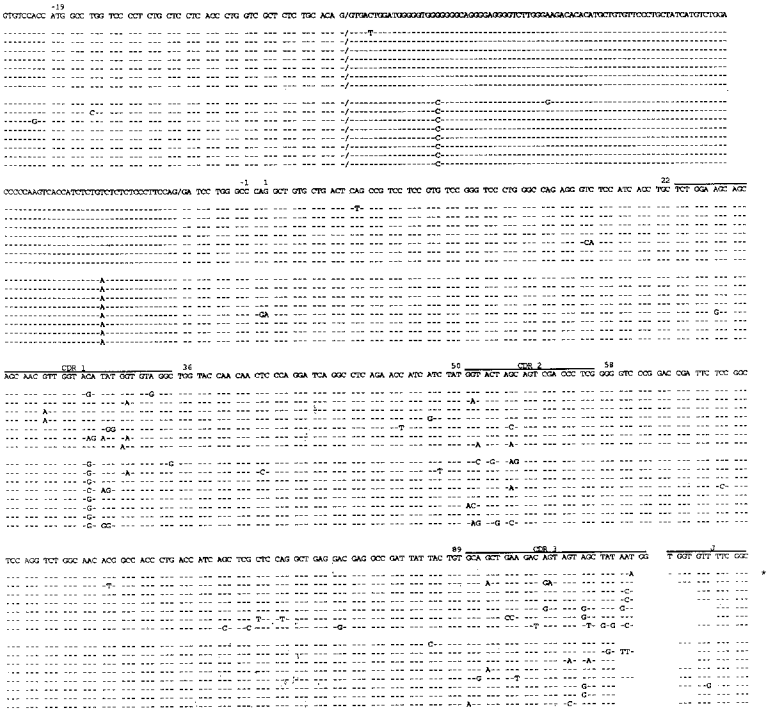

Figure 3. Somatic Mutations in Rearranged Vλ Sequences from IPP of Germ-Free Sheep

Rearrangement type II (Reynaud et al., 1991) was isolated after gel fractionation of EcoRI-digested IPP DNA from a germ-free sheep at 5 weeks of age and was cloned after PCR amplification. Sequences are compared with the corresponding gene, Vλ 16.1, and are distributed in two allelic forms, one identical to 16.1 and the other one differing at three positions. The asterisk indicates an out-of-frame V–J junction.

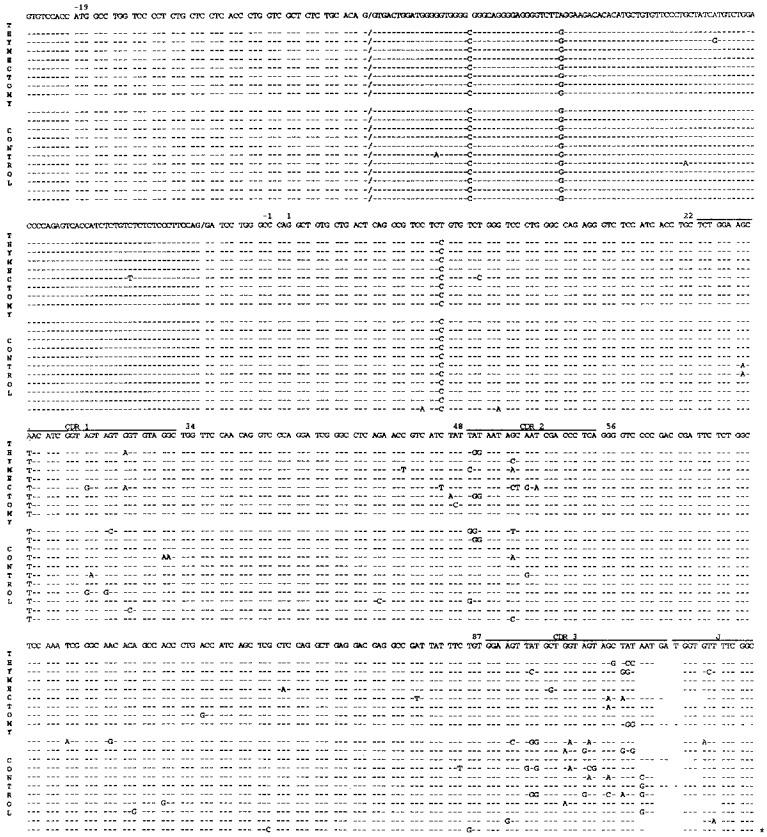

Figure 4. Somatic Mutations in Rearranged Vλ Sequences Isolated from IPP of a Thymectomized Sheep

Rearranged Vλ sequences were obtained as described in Figure 3 from a 5-week-old sheep thymectomized at day 70 in utero (THYMECTOMY) and from its unoperated twin (CONTROL) and were compared with the 5.1 gene. The asterisk indicates an out-of-frame V–J junction.

this could have any effect on the hypermutation process taking place in IPP B cells.

Phenotypic Analysis

Peripheral lymphocytes from this animal were analyzed at three intervals after birth (4 days, 19 days, 33 days) and confirmed that its blood contained on average around 10% of normal $CD4^+$ and $CD8^+$ T cells and less than 1% of normal γδ T cell numbers, consistent with earlier results (Hein et al., 1990). This degree of peripheral T cell depletion had no obvious effect on the size of IPP follicles (Figure 1) and did not cause a discernable reduction in the already low number of $CD4^+$ and $CD8^+$ T cells normally found in the interfollicular areas of the IPP. However, there were many fewer γδ T cells in the intestinal mucosa of this animal. Fetal thymectomy had no apparent effect on the expression of B cell markers or the numbers of tingible body macrophages and dendritic cells.

Molecular Studies

Rearranged VJ fragments I (5.1) and II (16.1) were isolated and sequenced as in previous experiments from IPP DNA from thymectomized and control lambs sacrificed on the same day (day 33). The IPP VJ sequences from the thymectomized animal showed modifications similar to the control with a clustering of mutations in CDRs and biased R/S ratios (Figure 4 and Table 1).

A Large Vλ Genomic Pool

In our previous study, we isolated 18 different Vλ genes from a genomic EcoRI library. We wanted to find out what the total size of the sheep Vλ genomic pool could be and whether we could identify putative donors for the somatic modifications observed. We therefore constructed a partial Sau3A DNA library from another animal, from which six different Vλ-hybridizing phages were isolated and 11 Vλ genes sequenced (data not shown). Among these 11 genes, 2 were identical in their coding region to the previously isolated Vλ genes 18 and 26.2 (this last one being a pseudogene). Seven Vλ genes showed 5–27 differences over the V coding part, with the closest related Vλ gene among those previously described. The last two genes, with more than 60 bp differences from the 5.1 gene, were rather different from all the genes of the previous pool. Among this new set, there were three pseudogenes that present frameshift deletions in the coding region. One (out of six for which complete signals were determined) showed variant recombination signal sequences that may reduce its rearrangement frequency. The overall proportion of pseudogenes and of genes with inefficient rearrangement signals is about 1/5 and 1/3, respectively, in the total V pool.

Based on this new analysis coming from a different lamb with only 2 Vλ genes out of 11 being strictly identical to the previous 18, we have to reevaluate our previous estimate of the size of the sheep Vλ genomic pool, which we think must be composed of approximately 90–100 Vλ genes. This figure is, however, an upper limit, since new V genes with few differences from the previous ones could represent allelic variants.

Table 2. Nucleotide Substitutions in Sheep IPP V Sequences[a]

to:	G	A	T	C	Total	
from:						
G	–	**25.3%**	**4.3%**	**8.6%**	**38.2%**	
	–	29.3%	3.5%	7.8%	40.6%	**68.6%**
						69.1%
A	**22.5%**	–	**4.0%**	**3.8%**	**30.4%**	
	19.5%	–	5.1%	3.9%	28.5%	
T	**8.4%**	**7.3%**	–	**5.6%**	**21.3%**	
	7.0%	7.0%	–	7.8%	21.9%	**31.4%**
						30.9%
C	**2.6%**	**1.4%**	**6.1%**	–	**10.1%**	
	2.0%	2.0%	5.0%	–	9.0%	
Transitions	**59.5%**	**G/A**	**47.8%**	**T/C**	**11.7%**	
	61.6%		48.8%		12.8%	
Transversions	**40.5%**					
	39.1%					

[a] The mutations listed in Table 1 are enclosed. Figures in bold represent all 580 mutations listed, figures in plain type represent the mutations obtained in the loop (6 weeks and 8 weeks) and germ-free (7L and 10L) experiments (260 mutations). The nucleotide composition of the 5.1 and the 16.1 genes is as follows: 46% A/G, 54% C/T and 47% A/G, 53% C/T, respectively.

Nevertheless, with these additional data, we could not find any evidence for the presence of genomic donors for the somatic modifications observed in the previous and actual analysis.

Specificity of the Hypermutation Process

For 110 somatic VJ sequences analyzed in this study, there are 580 mutations, with 432 in CDRs and 95 in FRs, 24 in the J segment and 24 in the leader intron (Table 1).

The modifications are mainly single point mutations, but they also occur in doublet. Doublets appear early in the process (two in the fetal genomic sequences). There are only three cases of insertion or deletion, and this low number could have been created by the Taq polymerase as well. Out-of-frame V–J junctions are rare (3 out of 110 sequences). We proposed previously that, as described for chicken B cells, rearrangement would be attempted only on one allele due to a short window of time during which rearrangement would be performed, cells with unproductive rearrangements being discarded (Reynaud et al., 1991).

There is a strong bias for G↔A transitions, which amount to about half of all mutations (Table 2). Overall, there are 2.2 more mutations on purines than on pyrimidines, mutations being counted arbitrarily on the sense strand of DNA. This trend for purines and transitions over transversions was identical in the normal and the antigen-free situations (Table 2). This bias was confirmed on a large sample of L–V intronic sequences, a DNA segment that is not submitted to antigenic selection. A total of 50 intronic mutations were collected from both the present and the previous studies, as well as from an additional set of 45 intron sequences from 4 month IPP DNA (data not shown). After subtraction of the background of Taq polymerase errors, the same 2.2 ratio of purine versus pyrimidine mutations was obtained.

A Subset of CDR Codons Concentrates 90%–95% of All CDR Mutations

The mutations obtained in this study (351 mutations; Table 1), together with the "adult" sequences reported previously (170 mutations; Reynaud et al., 1991), were plotted along the 5.1 V sequence (Figure 5a), with differentiation of normal (405 mutations) versus loop conditions (116 mutations). For the 16.1 gene, an additional set of 25 control sequences at 5 weeks was obtained (data not shown), providing a total of 163 mutations in normal conditions versus 81 in germ-free conditions (Figure 5b).

When CDRs of the 5.1 gene are analyzed, 95% of all mutations are clustered in 17 amino acids out of 27 (positions with more than seven mutations). For the 16.1 gene, 16 codons out of 29 concentrate 92% of CDR mutations (positions with more than three mutations). This does not appear to result from structural constraints on specific amino acids of the CDRs, since, apart from one or two positions (e.g., Val-32 in the CDR1 of the 5.1 gene), silent mutations are not specifically favored at the positions with few mutations. The restricted set of mutated positions are very similar in the normal and the loop or germ-free conditions, suggesting that it represents a targeting at the DNA level, rather than specific spots selected by external antigens. Some differences exist (e.g., codon 90 in the 16.1 gene), but the sample size of the germ-free mutations may be too small to consider them significant.

Clustering of mutations only in the CDR1 of a mouse Vκ passenger transgene and the occurrence of specific "hot spots" of mutation have also led Betz et al. (1993a) to suggest a specific DNA targeting in the mouse hypermutation mechanism. We also observed among the "CDR-mutated subset" such highly mutated positions (e.g., GGT in CDR1; TAT, AGC in CDR2; AGT, AGC, TAT in CDR3 of the 5.1 gene, each one having more than 30 mutations out of a total of 408 in CDRs). Some, but not all, of these positions fit the consensus hot spot sequence found in

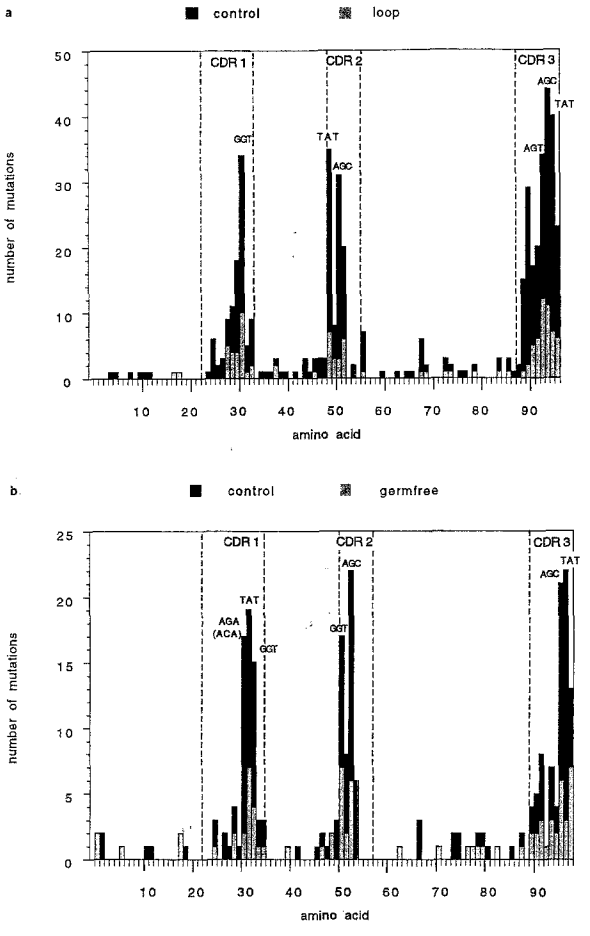

Figure 5. Pattern of Mutations of the 5.1 and 16.1 Genes in Loop, Germ-Free, and Normal Conditions

(a) Mutations obtained for the 5.1 V coding sequence and reported in Table 1 are represented, with discrimination of the mutations from the loop experiment (stippled areas, 116 mutations) versus all the other ones as "control" (closed areas, 405 mutations). The control sample includes the mutations of the adult (i.e., 4-month-old) IPP sample reported previously (Reynaud et al., 1991). The ordinate represents the number of mutations at each amino acid position. Limits of CDRs are delineated. The particular amino acid positions with 30 mutations or more (CDR hot spots) are indicated.

(b) Mutations obtained for the 16.1 coding sequence and reported in Table 1 are represented, with discrimination of the mutations from the germ-free experiment (stippled areas, 81 mutations) versus all the other ones as control (closed areas, 163 mutations). An additional set of 25 sequences (representing a total of 104 mutations) was determined for this histogram to reach a sufficient sample size and included in the control sample. The particular amino acid positions with 15 mutations or more (CDR hot spots) are indicated.

mouse sequences (RGYW, with R = G or A, Y = C or T, and W = A or T; Rogozin and Kolchanov, 1992; Betz et al., 1993b).

The CDRs of the Most Utilized Genomic Vλ Genes Have a Specific Base Pair Composition

The analysis of the Vλ germline gene of rearrangements I and II, which represent approximately 50% of the rearrangement events, shows a striking bias for G or A in the first and second position of each CDR codon (75% and 77%, respectively) and a bias against G or A in the third base (18% and 23%) (Table 3). Most importantly, this representation is still more biased in the subset of CDR residues in which over 90%–95% of all mutations are clustered: in this CDR-mutated subset, the first two positions are 88% and 81% G/A in the 5.1 and 16.1 gene, respectively, whereas 0% and 19% G/A are found at the third position of these genes.

The other functional Vλ genomic genes, although we do not know their exact usage frequency, show the same trend for G or A in the first two bases of their CDR codons, with some variation. For instance, for gene 12.2 (Reynaud et al., 1991), 76% of G/A are found in the first two bases of triplets in CDRs and 17% of G/A in the third base. For gene 18, which has been isolated from the two different libraries, there are 61% of G/A in the first two positions and 34% of G/A in position 3 for CDR triplets.

The Hypermutation Mechanism Targets Replacement Substitutions to CDRs

Two figures of the somatic mutation pattern have suggested to us that a strong antigenic selection occurs while mutant B cells proliferate in sheep IPP: first, a high clustering of mutation in CDRs (resulting in a 15-fold higher mutation frequency per nucleotide in CDRs versus FRs at the stage studied, i.e., 4 months of age); second, an R/S ratio of nucleotide replacement higher than the theoretical value in the CDRs (8.1 for a theoretical value of 3.6). In the present study, these values are 11.4 for the clustering of mutation and 11.0 for the R/S ratio (Table 1) and, most importantly, are the same in the different conditions studied.

However, clustering of mutations in CDRs appears to be created at least in part by the mechanism itself: when comparing silent mutations only, a 2.7-fold clustering in CDRs versus FRs is observed per nucleotide (36 and 30 silent mutations in CDRs and FRs, respectively, for a length ratio of 2.56; cf. Table 1). When translated into total mutations, i.e., R plus S mutations, two corrections must be made: first, about one third of the mutations are "lost" from FR sequences due to structural constraints on the IgM molecule (R/S of 1.8 instead of the theoretical 2.9). Another type of correction must be made for CDRs: we have calculated a "mechanistic" R/S value, which differs from the theoretical one, by taking into account the bias for GA over CT replacement (see Experimental Procedures for a detailed calculation); this gives a 4.7 R/S estimate, higher than the 3.6 theoretical one, because CDR codons have a high GA content at the first two nucleotide positions and a high CT content at the third one. One can thus estimate a "mechanistic plus structural" clustering, i.e., without invoking antigen selection, by the ratio of

$$\frac{(R+S)CDR}{(R+S)FR} \times \text{length ratio}\ \frac{FR}{CDR}, \text{ or } \frac{(R/S+1)CDR}{(R/S+1)FR} \times \text{silent mutation clustering}\ \frac{CDR}{FR}, \text{ i.e. } \frac{(4.7+1)}{(1.8+1)} \times 2.7 \text{ of } 5.5.$$

The second point concerns the R/S ratio. As mentioned previously, a mechanistic R/S of 4.7 is estimated for CDRs.

Table 3. Percentage of Purines in Codons of Sheep V λ 5.1 and 16.1 Genes

	FR		CDR Total		CDR Mutated Subset[a]	
Nucleotide Position in the Codon	First + Second	Third	First + Second	Third	First + Second	Third
V 5.1	47%	39%	75%	18%	88%	0%
V 16.1	48%	39%	77%	23%	81%	19%

[a] CDR-mutated subsets are defined as amino acid postions in CDR with more than seven mutations for the 5.1 gene and more than three mutations for the 16.1 gene (as listed in Figures 6 and 7). For 5.1, it represents 17 amino acids out of 27, accounting for 95% of all CDR mutations (CDR1, positions 28–31, 33; CDR2, 49–52, CDR3, 89–96). For 16.1, it represents 16 amino acids out of 29, accounting for 92% of total CDR mutations (CDR1, positions 29, 31–35; CDR2, 51–54; CDR3, 90–92, 94–98).

However, the effective value can be estimated to be much higher. Of the total mutations, 90%–95% are indeed clustered in half of the CDR1 and CDR2, and most of the CDR3, in what we name the CDR-mutated subset. When only these positions are considered for estimating the R/S ratio, and again taking into account the 2-fold higher mutation rate of GA over CT nucleotides, a "selection-free" R/S value of 8.7 is then obtained for the 5.1 gene, and 5.9 for the 16.1 (see calculation in Experimental Procedures). The observed values in the previous and the present experiments are, respectively, 8.2 and 13.8 for the 5.1 gene, and 7.7 for the 16.1.

These calculations are only indicative, since the estimate of silent mutations in CDRs is probably imprecise; e.g., mutations can superimpose, or there could be false attributions in the case of mutations occurring in doublets.

Discussion

IPPs are the major primary site of B cell production in sheep during development (Reynolds and Morris, 1983a). 3×10^9 surface IgM B cells are produced per hour among which 5% migrate out of the organ, the majority of cells dying in situ (Reynolds, 1986). We showed that, while proliferating in IPP follicles, B cells diversify their antigen receptor by an ongoing hypermutation process operating at a rate comparable to what has been observed during antigen-induced immune responses in the mouse (0.5×10^{-4}–1.2×10^{-4}/bp per division). Moreover, replacement mutations accumulated over time in CDRs, implying, with the high cellular death rate observed in the organ, that a strong selective pressure was probably applied to the mutants (Reynaud et al., 1991).

We wanted to find out whether this positive selection was due to a signal given through the antigen receptor by internal ligands or external antigens. We therefore studied the level of hypermutation and the accumulation of replacement substitutions in CDRs of rearranged Vλ genes in experimental situations in which IPPs were isolated from foreign antigens during development. We first studied 2-month-old B cells from sterile segments of IPP that had been detached and isolated from the intestine during fetal life and therefore had never been in contact with gut-associated food and bacterial antigens. The level of mutation was identical when comparing the B cells from these ileal loops with those from the adjacent normal tissue, and replacement substitutions in the CDRs were similarly favored.

The same results were obtained when analyzing IPP B cells from germ-free sheep fed a sterilized diet. However, such gnotobiotic animals are still contaminated with food antigens, which represent a small proportion of the normal antigenic environment. The fact that these gnotobiotic animals do not show, like in the isolated IPP segments, any quantitative difference either in the number of mutations or in the CDR R/S ratios when compared with the normal situation implies that the hypermutation process is not influenced by foreign antigens. Animals thymectomized at 70 days of gestation, which are largely T cell depleted in the peripheral blood (Hein et al., 1990), also showed no difference with the control situation.

A 10-fold clustering of mutations in CDRs versus FRs, comparable to controls, is observed in the different experimental conditions. Mutations were located on specific codons of each CDR, and these mutated CDR subsets were very similar in the normal, loop, and germ-free conditions, thus arguing against a foreign antigen–driven selection process. Mutational hot spots that are selected by antigens during immune responses in the mouse can be clearly distinguished from the ones present on unselected passenger transgenes and thus introduced by the mechanism per se (Betz et al., 1993b).

The analysis by genomic cloning of fetal rearranged Vλ sequences confirmed, moreover, the results previously obtained by PCR amplification that there was on the average one modification per sequence at the time of birth. Collectively, these results show that the hypermutation mechanism taking place in IPP B cells is part of a developmentally regulated antigen-independent process that starts in the sterile environment of the fetus and continues for several months after birth. They also confirm that germinal centers induced by immunization with T-dependent antigens are not the only appropriate environment for somatic hypermutation of immunoglobulin genes. Although IPP follicles contain 95% of surface IgM B cells, they also contain a variety of other cellular components such as $CD4^+$ T cells, dendritic cells, tingible body macrophages, stromal cells, and reticular cells (Hein et al., 1989). Whether these different constituents play a role in triggering the diversification process remains unknown at the moment.

As reported before (Reynolds, 1976; Reynolds and Mor-

ris, 1983b), the ileal segment isolated from the gut showed a clear reduction in the size of the follicles at 1 month after birth. After 4 months, these isolated follicles have totally regressed and are filled with collagen (Reynolds, 1976). Such a reduction in the size of the follicles was also observed in the germ-free animals, suggesting that microbial components such as lipopolysaccharide must exert a mitotic stimulatory activity on follicular B cells. A similar inhibition of growth of gut-associated lymphoid follicles was previously reported in the chicken bursa isolated from gut constituents by embryonic bursal duct ligation (Ekino, 1993).

Mutations accumulate over time during development in IPP B cells, showing 0.9 modifications per sequence at 140 days of fetal life, 4–5 at 5–6 weeks, 7–8 at 8 weeks, 11 at 4 months, and up to 15 in peripheral B cells at 1.5 years of age (unpublished data). These numbers suggest that the process slows down at around 2 months and is probably minimal after 6 months of age, the IPP follicles starting to regress at this stage until their complete involution, which occurs around 1 year of age.

Having excluded the role of external antigens, what could be the selection processes acting on sheep IPP B cells and could the selective proliferation of the mutants be driven entirely by internal ligands? When looking at the VJ sequences we have analyzed in this and the previous study, there were very few deleterious mutations (as also monitored by the low R/S value of mutations in framework regions), implying that an IgM molecule has to be assembled and present at the cell surface for IPP B cells to survive. Moreover and surprisingly, when looking at the various components of this model, it becomes clear that the pattern of mutation could be due in some part to the intrinsic specificity of the process. Hypermutation does not occur at random in sheep IPP B cells. When looking at the sense strand of DNA, there was a strong bias for G↔A transitions (48% of all mutations), and, taking into account all types of mutations, they occurred 2.2 times more frequently on purines than on pyrimidines: this ratio is the same whether mutations on coding or noncoding (leader intron) sequences are considered. Moreover, the CDRs of the two major Vλ rearranged genes contain mostly G or A in the first two positions of their triplets (75%–77%) and a much lower proportion of G or A in the third base (18%–23%), this trend being even stronger in the subset of CDR codons in which most mutations are clustered. Consequently, when the bias toward G/A mutations is taken into account to calculate the theoretical R/S value of CDR codons, this value becomes very close to the one observed experimentally. Moreover, based on the repartition of silent mutations that show a 2.7-fold clustering in CDRs, and after extrapolation to total mutations (see Results), we estimated that the process itself generates as a minimum value 5.5 more mutations in CDRs than in FRs in the absence of antigen selection. Therefore, taking advantage of the G/A nucleotide specificity and of the bias in codon usage of the Vλ genes, the mechanism per se preferentially generates replacement mutations in CDRs in sheep IPP B cells.

A similar bias for mutations on purines and for concentration in CDRs was also observed on passenger immunoglobulin transgenes that are mutated but not submitted to antigen-induced selection during secondary immune responses in the mouse (Betz et al., 1993a), which suggests that the molecular mechanism is probably identical in both species. Such an "unselected" pattern reflects, therefore, the DNA specificity of the process. When mouse immunoglobulin V genes undergoing mutation are submitted to antigenic selection, some specific antigen-driven mutations are superimposed on this basic mechanistic pattern, and they may sometimes obscure the molecular trend. As noted in the mouse, the imbalance between purines and pyrimidine mutations indicates that the mutation mechanism is targeted to one strand of DNA. Models in which hypermutation could be achieved by error-prone DNA synthesis have been put forward (Brenner and Milstein, 1966), but there are no experimental data at the moment that can support any particular proposition.

In conclusion, in a model in which positive selection by external antigens can be assessed, we have shown that the specific pattern of mutations remains unchanged in the absence of foreign antigens and that the mechanism per se can preferentially introduce replacement mutations in CDRs. Nevertheless, it certainly remains difficult when estimating the R/S ratio and the clustering of mutation in CDRs to evaluate, at this level of approximation, the precise role of internal ligands that could drive some mutant B cells as they are produced. Moreover, whether foreign antigens in the normal situation will influence the distribution of the repertoire within IPP follicles or as B cells migrate out of the IPP to the periphery can not be concluded in this model.

There are different cellular and molecular strategies to generate an antibody repertoire, but overall, these strategies have evolved toward the same goal: the production of a large number of different antibody molecules carrying different antigen-recognizing domains. Why would certain species select for GALT or bone marrow to generate their B cell compartment? Our results do not point to a role of external antigens on the specificity of the B cell repertoire in animals using GALT. Molecular strategies have probably evolved along with the structure of primary B cell organs. In such a sense, the potent nonspecific mitotic activity provided by gut bacterial antigens could be the necessary stimuli allowing B cells to proliferate for several months and thus to accumulate by gene conversion or hypermutation modifications on their immunoglobulin receptors.

Experimental Procedures

Animals and Surgical Procedures

White Alpine sheep were obtained from Versuchsbetrieb Sennweid (Olsberg, Switzerland).

Isolation of Ileal Fragments

Segments of ileum were isolated surgically in six fetal lambs aged from 117–126 days of gestation, using described procedures (Reynolds and Morris, 1983b). The isolated segments of fetal gut were 8–10 cm long and commenced at a position about 5 cm cranial to the ileocaecal junction. Care was taken during surgery to preserve the normal vascular, lymphatic, and nerve supply to the isolated gut segments. Two operated fetuses were recovered on days 143 and 144 of gestation

(term is 150 days). The remainder were born naturally, maintained under normal husbandry conditions, and killed at intervals after birth. Consistent with previous experience (Reynolds, 1976; Reynolds and Morris, 1983b), isolated gut segments grew considerably after birth and all appeared healthy at necropsy. Postnatally, they became slightly distended at their caudal end, presumably due to peristaltic accumulation of shed epithelia. Samples of Peyer's patches were removed from both normal ileum and the nondistended part of the isolated "loops" for immunohistology and molecular analysis. In one animal, aged 101 days, lumenal contents were removed aseptically from both normal and isolated ileum for microbiological analysis.

Thymectomy

One fetal lamb was thymectomized on day 70 of gestation (Hein et al., 1990) and sampled as described above on day 33 after birth.

Germ-Free Sheep

Gnotobiotic lambs were prepared and maintained by Dr. L. H. Thomas, Institute for Animal Health (Compton, England), using established procedures. Lambs were derived from the Compton flock of polled Dorset ewes, delivered by hysterotomy through an aseptic lock into sterile isolators, and reared on a diet of sterilized reconstituted milk replacement powder. Solid food in the form of irradiated stock pellets was offered from 2.5 weeks of age. Swabs were taken aseptically twice-weekly from the rectum and nostrils of each lamb and cultured for standard microbiological analysis.

Immunohistology and FACS Analysis

The expression of lymphocyte differentiation markers was assessed by flow cytometry of freshly isolated IPP lymphocytes and by immunoperoxidase staining of frozen sections of IPP derived from both normal and experimental animals. Monoclonal antibodies (MAbs) specific for the sheep CD5, CD4, CD8 γδ TCR and CD45RA molecules as well as undefined determinants expresssed specifically on B cells (MAbs Du2-8 and Du2-54) were used for FACS analysis. Additional antibodies specific for dendritic cells (MAb V-5) and tingible body macrophages (MAb Du2-66) were used for immunohistology. The procedures used during staining were as described (Hein et al., 1989).

DNA Isolation, DNA Amplification, and Sequencing

DNA preparation, isolation of the size region of rearrangement type I and type II after fractionation by gel electrophoresis, PCR amplification of VJ rearranged sequences, cloning in M13mp18, and sequencing were done as described previously (Reynaud et al., 1991). On the average, 1 out of 10 sequences from the size region of rearrangement type I was either a different gene or a PCR artifact resulting in a hybrid sequence, the other ones corresponding to the same germline gene (with four differences from the Vλ 5.1 gene). For rearrangement type II, variations were observed between different animals. In the germ-free sheep 7L, most sequences were identical to the 16.1 gene. In another animal (germ-free 10L), the new 16/5 gene dominated. In the thymectomy experiment, about half of the sequences only corresponded to the 16.1 gene, the other ones coming from various genes or being hybrids.

Genomic Cloning

To obtained fetal rearranged genomic clones, EcoRI-digested DNA from fetal IPPs (day 144 of gestation) was cloned in EMBL4 after size fractionation on sucrose gradients and selection of DNA in the 20 kb range. Additional germline Vλ sequences were obtained from a partial Sau3A DNA library from fetal liver, originating from the same animal from which fetal IPP rearranged sequences have been isolated (Reynaud et al., 1991)

Calculations of R/S Ratios

Theoretical R/S ratios were calculated for CDRs of 5.1 and 16.1 genes by adding replacement and silent substitutions of each amino acid. A mechanistic R/S ratio was calculated by substituting a 2.2 factor for each mutation bearing on a G or an A nucleotide. For example, the AGT codon has an R/S value of 8 (8 replacement mutations for 1 silent mutation); if G/A are mutated 2.2 more frequently than C/T, then there are (6 × 2.2 + 2 × 1) replacement mutations (6 replacement mutations on a G or an A nucleotide and 2 on a T) for 1 silent mutation (on a T), i.e., an R/S of 15.2. The values obtained for CDRs are as follows: for the gene 5.1, the theoretical value is 3.6, and the mechanistic one is 4.7; for 16.1, the two values are 3.4 and 4.5. When only the CDR-mutated subsets are considered (see Results), mechanistic R/S ratios are 8.7 for the 5.1 gene and 5.9 (or 5.4 depending upon the allelic form) for the 16.1 gene.

Acknowledgments

We thank Lisbeth Dudler, Birgit Kugelberg, and Annie de Smet for skilled technical assistance, Frederic Marin for the analysis of genomic V genes, Dr. I. Bloch (Veterinäramt Kanton Basel-Stadt) for performing the microbiological testing of gut lumen contents, Louis Du Pasquier for critical reading of the manuscript, and Charley Steinberg for strong, but constructive criticism. This work was supported by grants from the Ministère de la Recherche Scientifique and the Fondation pour la Recherche Médicale. We thank the Association Française contre les Myopathies for general support in the laboratory installation in Paris. The Basel Institute for Immunology was founded and is supported by F. Hoffmann-La Roche Ltd. (Basel, Switzerland).

Received September 5, 1994; revised October 10, 1994.

References

Becker, R. S., and Knight, K. L. (1990). Somatic diversification of immunoglobulin heavy chain VDJ genes: evidence for somatic gene conversion in rabbits. Cell *63*, 987–997.

Betz, A. G., Rada, C., Pannell, R., Milstein, C., and Neuberger, M. S. (1993a). Passenger transgenes reveal intrinsic specificity of the antibody hypermutation mechanism: clustering, polarity, and specific hot spots. Proc. Natl. Acad. Sci. USA *90*, 2385–2388.

Betz, A. G., Neuberger, M., and Milstein, C. (1993b). Discriminating intrinsic and antigen-selected mutational hotspots in immunoglobulin V genes. Immunol. Today *14*, 405–411.

Bos, N. A., and Meeuwsen, C. G. (1989). B cell repertoire in adult antigen-free and conventional neonatal BALB/c mice. I. Preferential utilization of the C_H–proximal V_H gene family PC7183. Eur. J. Immunol. *19*, 1811–1815.

Brenner, S., and Milstein, C. (1966). Origin of antibody variation. Nature *211*, 242–243.

Deenen, G. J., and Kroese, F. G. M. (1993). Kinetics of B cell subpopulations in peripheral lymphoid tissues: evidence for the presence of phenotypically distinct short-lived and long-lived B cell subsets. Int. Immunol. *5*, 735–741.

Ekino, S. (1993). Role of environmental antigens in B cell proliferation in the bursa of Fabricius at neonatal stage. Eur. J. Immunol. *23*, 772–775.

Goodnow, C. C., Crosbie, J., Adelstein, S., Lavoie, T. B., Smith-Gill, S. J., Brink, R. A., Pritchard-Briscoe, H., Wotherspoon, J. S., Loblay, R. H., Raphael, K., Trent, R. J., and Basten, A. (1988). Altered immunoglobulin expression and functional silencing of self-reactive B lymphocytes in transgenic mice. Nature *334*, 676–682.

Gu, H., Tarlinton, D., Müller, W., Rajewsky, K., and Förster, I. (1991). Most peripheral B cells in mice are ligand selected. J. Exp. Med. *173*, 1357–1371.

Hein, W. R., Dudler, L., and Mackay, C. R. (1989). Surface expression of differentiation antigens on lymphocytes in the ileal and jejunal Peyer's patches of lambs. Immunology *68*, 365–370.

Hein, W. R., Dudler, L., and Morris, B. (1990). Differential peripheral expansion and *in vivo* antigen reactivity of α/β and γ/δ T cells emigrating from the early fetal lamb. Eur. J. Immunol. *20*, 1805–1813.

Huetz, F., Carlsson, L., Tornberg U.-C., and Holmberg, D. (1993). V-region directed selection in differentiating B lymphocytes. EMBO J. *12*, 1819–1826.

Lafaille, J. J., DeCloux, A., Bonneville, M., Takagaki, Y., and Tonegawa, S. (1989). Junctional sequences of T cell receptor γδ genes: implications for γδ T cell lineages and for a novel intermediate of V-(D)-J joining. Cell *59*, 859–870.

Landsverk, T. (1987). The follicle-associated ephilium of the ileal Peyer's patch in ruminants is distinguished by its shedding of 50 nm particles. Immunol. Cell Biol. *65*, 251–261.

Lawton, A. R., and Cooper, M. D. (1974). Modification of B lymphocyte differentiation by anti-immunoglobulins. Contemp. Top. Immunobiol. *3*, 193–225.

Lederberg, J. (1959). Genes and antibodies. Science *129*, 1649–1653.

Ma, A., Fisher, P., Dildrop, R., Oltz, E., Rathbun, G., Achacaso, P., Stall, A., and Alt, F. W. (1992). Surface IgM mediated regulation of RAG gene expression in Eμ-N-myc B cell lines. EMBO J. *11*, 2727–2734.

Metcalf, E. S., and Klinman, N. R. (1976). *In vitro* tolerance induction of neonatal murine B cells. J. Exp. Med. *143*, 1327–1340.

Nemazee, D. A., and Bürki, K. (1989). Clonal deletion of B lymphocytes in a transgenic mouse bearing anti-MHC class I antibody genes. Nature *337*, 562–566.

Nossal, G. J. V., and Pike, B. L. (1975). Evidence for the clonal abortion theory of B lymphocyte tolerance. J. Exp. Med. *141*, 904–917.

Reynaud, C.-A., McKay, C. R., Müller, R. G., and Weill, J.-C. (1991). Somatic generation of diversity in a mammalian primary lymphoid organ: the sheep ileal Peyer's patches. Cell *64*, 995–1005.

Reynolds, J. D. (1976). The development and physiology of the gut-associated lymphoid system in lambs. PhD thesis, Australian National University, Canberra, Australia.

Reynolds, J. D. (1986). Evidence of extensive lymphocyte death in sheep Peyer's patches. I. A comparison of lymphocyte production and export. J. Immunol. *136*, 2005–2010.

Reynolds, J. D., and Morris, B. (1983a). The evolution and involution of Peyer's patches in fetal and postnatal sheep. Eur. J. Immunol. *13*, 627–635.

Reynolds, J. D., and Morris, B. (1983b). The influence of gut function on lymphoid cell populations in the intestinal mucosa of lambs. Immunology *49*, 501–509.

Rogozin, I. B., and Kolchanov, N. A. (1992). Somatic hypermutagenesis in immunoglobulin genes. II. Influence of neighbouring base sequences on mutagenesis. Biochim. Biophys. Acta *1171*, 11–18

Rolink, A., and Melchers, F. (1991). Molecular and cellular origins of B lymphocyte diversity. Cell *66*, 1081–1094.

Viale, A.-C., Coutinho, A., and Freitas, A. A. (1992). Differential expression of V_H families in peripheral B cell repertoires of newborn or adult immunoglobulin H chain congenic mice. J. Exp. Med. *175*, 1449–1456.